ERROR BOUNDS FOR MONOTONE APPROXIMATION SCHEMES FOR HAMILTON-JACOBI-BELLMAN EQUATIONS

GUY BARLES AND ESPEN R. JAKOBSEN

Abstract. We obtain error bounds for monotone approximation schemes of Hamilton-Jacobi-Bellman equations. These bounds improve previous results of Krylov and the authors. The key step in the proof of these new estimates is the introduction of a switching system which allows the construction of approximate, (almost) smooth supersolutions for the Hamilton-Jacobi-Bellman equation.

1. Introduction

This paper is a continuation of a work started in [2] (see also Jakobsen [21]) whose aim is to prove results on the rate of convergence of monotone approximation schemes for possibly degenerate Hamilton-Jacobi-Bellman equations (HJB equations in short) by purely analytical methods. Krylov [26, 27] obtained such results in a rather general framework but by using a combination of PDE arguments and rather deep probabilistic estimates which we want to avoid.

The strategy we used in [2] is based on the idea that the HJB equation and the approximation scheme should play symmetrical roles. Unfortunately, this leads to unnatural restrictions on the data when the scheme in consideration is a finite difference method. These restrictions do not appear in [27]. In the present paper, we use a more classical strategy in which the HJB equation plays the central role. Our approach yields results in the full generality, improving those of [26, 27] and [2].

In order to be more specific, we introduce the HJB equation which is written in the form

$$F(x,u,Du,D^2u) = 0 \quad \text{in} \quad \mathbb{R}^N, \tag{1.1}$$

with

$$F(x,t,p,X) = \sup_{\alpha\in\mathcal{A}} \mathcal{L}^\alpha(x,t,p,X),$$
$$\mathcal{L}^\alpha(x,t,p,X) = -\mathrm{tr}[a^\alpha(x)X] - b^\alpha(x)p + c^\alpha(x)t - f^\alpha(x),$$

Date: January 5, 2004.

Key words and phrases. Hamilton-Jacobi-Bellman equations, switching system, viscosity solution, approximation schemes, finite difference methods, convergence rate, error bound.

where tr denotes the trace. The coefficients a, b, c, f are, at least, continuous functions defined on $\mathbb{R}^N \times \mathcal{A}$ with values respectively in the space $S(N)$ of symmetric $N \times N$ matrices, $\mathbb{R}^N$ and $\mathbb{R}$. The space of controls, $\mathcal{A}$, is assumed to be a compact metric space. Precise assumptions on the data will be given later on. Under classical assumptions, it is well-known that this equation is associated to a stochastic optimal control problem, and that the value function of this problem, is the unique viscosity solution of the equation. Moreover, the value function is typically bounded and Hölder continuous, and the regularity depends on the properties of a, b, c and f.

The monotone approximation schemes we consider are of the following type,

$$S(h, x, u_h(x), [u_h]_x) = 0 \quad \text{in} \quad \mathbb{R}^N, \tag{1.2}$$

where S is, loosely speaking, a consistent, monotone and uniformly continuous approximation of F in (1.1). The approximate solution is u_h, $[u_h]_x$ is a function defined from u_h, and the approximation parameter is h. This abstract notation was introduced by Barles and Souganidis [3] to display clearly the monotonicity of the scheme: S is non-decreasing in u_h and non-increasing in $[u_h]_x$ with the classical ordering for functions. Typical approximation schemes which we have in mind, are finite difference methods (FDMs) and control schemes based on the dynamic programming principle. We refer to Dupuis and Kushner [11] and Camilli and Falcone [5] for more information about such schemes.

In the viscosity solutions setting the first results on convergence rates for monotone schemes were obtained by Crandall and Lions [10]. Later the first-order case have been studied by many authors considering different schemes and assumptions [7, 36, 37, 13, 1, 25, 29, 35, 28, 24]. Only recently did Krylov [26, 27] obtain the first results for second-order equations (for HJB equations), and these results were then partially extended by Barles and Jakobsen [2, 21]. These results concern only HJB equations, or equivalently, equations with convex/concave Lipschitz continuous non-linearity F. In the non-convex (or non-concave) case, to the best of our knowledge, there are no general results. There exist results only in particular cases like, for example, in one space-dimension [20] and for obstacle problems [19].

From a technical point of view, the upper estimate on $u - u_h$ is much easier to obtain than the lower estimate. Roughly speaking, a regularization of the solution u by convolution provides approximate smooth subsolutions of the equation because of the convexity of the equation. By inserting this smooth subsolution in the scheme and using consistency, one is led to the upper bound after choosing an optimal parameter of regularization. It is worth pointing out that a non-trivial difficulty in performing this argument is the x-dependence in the equation. This difficulty was solved by a very clever argument of Krylov [27] which is used extensively in [2] and in the present paper.

Unfortunately, this is clearly a one-sided argument working only for convex equations. In general, there is no simple way to build approximate smooth supersolutions which would lead to the lower estimate on $u - u_h$. It is precisely this difficulty that we overcome here. In fact, we do not really build a sequence of approximate smooth supersolutions, but a sequence of supersolutions which behave as if they were smooth. The key step here is to introduce switching system approximations of the HJB equation and study their rates of convergence. This approach is inspired by Evans and Friedman [12], see also [6]. These rates of convergence are obtained by combining the above mentioned idea of Krylov and an approach suggested by Lions [33]. Even if we do not make a point of proving general results in this direction, this part has an independent interest. It seems to be the first time that the rate of convergence is obtained for such switching system approximations in the case of second-order equations.

In order to give a flavor of our results, for HJB equation satisfying natural assumptions and with bounded Lipschitz continuous solutions, we prove a lower estimate of the form $h^{1/5}$ for a standard finite difference method. The corresponding result in Krylov [27] was $h^{1/27}$.

The paper is organized as follows: In Section 2 we introduce the switching system and prove the rate of convergence. This result is then used in Section 3 for obtaining the rate of convergence of the approximation scheme (1.2). In Section 4, we apply the result of Section 3 to a typical finite difference method for the HJB equation taken from Dupuis and Kushner [11]. In order to simplify the exposure, the proofs in the paper are presented in a context where all the solutions are Lipschitz continuous. In Section 5, we provide without proofs, extensions to the case of $C^{0,\delta}(\mathbb{R}^N)$-solutions. We also discuss the fact that our approach is rather close to provide results for the non-convex (non-concave) case. Finally the Appendix collects several results for switching systems (well-posedness, regularity, and continuous dependence) which are used throughout the paper.

We conclude this introduction by explaining the notation we will use throughout this paper. By $|\cdot|$ we mean the standard euclidian norm in any $\mathbb{R}^p$ type space (including the space of $N \times P$ matrices). In particular, if $X \in S(N)$ then $|X|^2 = \mathrm{tr}(XX^T)$ where X^T denotes the transpose of X. Now if w is a bounded function from $\mathbb{R}^N$ into either $\mathbb{R}$, $\mathbb{R}^M$, or the space of $N \times P$ matrices, we set

$$|w|_0 = \sup_{y \in \mathbb{R}^N} |w(y)|.$$

If w is also Lipschitz continuous, we set

$$[w]_1 = \sup_{x \neq y} \frac{|w(x) - w(y)|}{|x - y|} \quad \text{and} \quad |w|_1 = |w|_0 + [w]_1.$$

We denote by $\leq$ the component by component ordering in $\mathbb{R}^M$ and the ordering in the sense of positive semi-definite matrices in $S(N)$. For the rest of this paper we let ρ denotes the same, fixed, positive smooth function with support in $\{|x| < 1\}$ and mass 1. From this function ρ, we define the sequence of mollifiers $\{\rho_\varepsilon\}_{\varepsilon>0}$ as follows,

$$\rho_\varepsilon(x) = \frac{1}{\varepsilon^N}\rho\left(\frac{x}{\varepsilon}\right) \quad \text{in} \quad \mathbb{R}^N.$$

We also use the following spaces: $C_b(\mathbb{R}^N)$ and $C^{0,\delta}(\mathbb{R}^N)$, $\delta \in (0,1]$, denoting respectively the space of bounded continuous functions on $\mathbb{R}^N$ and the space of bounded δ-Hölder continuous functions on $\mathbb{R}^N$.

2. Convergence Rate for a Switching System.

In this section, we obtain the rate of convergence for certain switching system approximations to the HJB equation (1.1). Such approximations have be studied in [12, 6], and a viscosity solutions theory of switching systems can be found in [38, 18, 17]. We consider the following type of switching systems,

$$F_i(x, v, Dv_i, D^2v_i) = 0 \quad \text{in} \quad \mathbb{R}^N, \quad i \in \mathcal{I} := \{1, \dots, M\}, \tag{2.1}$$

where the solution $v = (v_1, \cdots, v_M)$ is in $\mathbb{R}^M$, and for $i \in \mathcal{I}$, $x \in \mathbb{R}^N$, $r = (r_1, \cdots, r_M) \in \mathbb{R}^M$, $p \in \mathbb{R}^N$, and $X \in \mathcal{S}^N$, F_i is given by

$$F_i(x, r, p, X) = \max\Big\{ \sup_{\alpha\in\mathcal{A}_i} \mathcal{L}^\alpha(x, r_i, p, X);\ r_i - \mathcal{M}_i r\Big\},$$

where $\mathcal{A}_i \subset \mathcal{A}$, $\mathcal{L}^\alpha$ is defined below (1.1), and for $k > 0$,

$$\mathcal{M}_i r = \min_{j\neq i}\{r_j + k\}.$$

Under suitable assumptions on the data, we have existence and uniqueness of a solution v of this system. Moreover, it is not difficult to see that, as $k \to 0$, every component of v converge locally uniformly to the solution of the following HJB equation

$$\sup_{\alpha\in\overline{\mathcal{A}}} \mathcal{L}^\alpha(x, u, Du, D^2u) = 0 \quad \text{in} \quad \mathbb{R}^N\ , \tag{2.2}$$

where $\overline{\mathcal{A}} = \cup_i \mathcal{A}_i$.

The objective of this section is to obtain an error bound for this convergence. For the sake of simplicity, we restrict ourselves to the situation where the solutions are Lipschitz continuous. However, it is not difficult to adapt our approach to more general situations, and we give results in this direction in Section 5.

We will use the following assumptions:

(A1) For any $\alpha \in \mathcal{A}$, $a^\alpha = \frac{1}{2}\sigma^\alpha\sigma^{\alpha T}$ for some $N \times P$ matrix σ^α. Furthermore, there are constants λ, K independent of α such that

$$c \geq \lambda > 0 \quad \text{and} \quad |\sigma^\alpha|_1 + |b^\alpha|_1 + |c^\alpha|_1 + |f^\alpha|_1 \leq K.$$

(A2) The constant λ in (A1) satisfies $\lambda > \sup_\alpha \left\{[\sigma^\alpha]_1^2 + [b^\alpha]_1\right\}$.

As the reader will see below and in the following sections, assumption **(A1)** ensures the well-posedness of all the equations and systems of equations we consider in this paper. If we assume in addition (A2), all solutions will belong to $C^{0,1}(\mathbb{R}^N)$. We refer to the Appendix for a precise justification of these claims. In the present situation, we have the following well-posedness and regularity result.

Proposition 2.1. *(i) Assume (A1). If w_1 and w_2 are sub- and supersolutions of* (2.1) *or* (2.2)*, then $w_1 \le w_2$.*

(ii) Assume (A1) and (A2). Then there exist unique solutions v and u of (2.1) *and* (2.2) *respectively, satisfying*

$$|v|_1 + |u|_1 \le C,$$

where the constant C only depends on K, λ from (A1).

In order to obtain the rate of convergence for the switching approximation, we use the before mentioned regularization procedure of Krylov [27, 2]. This procedure requires the introduction of following auxiliary system

$$F_i^\varepsilon(x, v^\varepsilon, Dv_i^\varepsilon, D^2 v_i^\varepsilon) = 0 \quad \text{in} \quad \mathbb{R}^N, \quad i \in \mathcal{I}, \tag{2.3}$$

where $v^\varepsilon = (v_1^\varepsilon, \cdots, v_M^\varepsilon)$,

$$F_i^\varepsilon(x, r, p, M) = \max\Big\{ \sup_{\alpha\in\mathcal{A}_i, |e|\le\varepsilon} \mathcal{L}^\alpha(x+e, r_i, p, X);\ r_i - \mathcal{M}_i r\Big\},$$

and $\mathcal{L}$ and $\mathcal{M}$ are defined below (1.1) and (2.1) respectively. By Theorems A.1 and A.3 in the Appendix, we have the following result:

Proposition 2.2. *(i) Assume (A1). If w_1 and w_2 are sub- and supersolutions of* (2.3)*, then $w_1 \le w_2$.*

(ii) Assume (A1) and (A2). Then there exist a unique solution v^ε of (2.3) *satisfying*

$$|v^\varepsilon|_1 + \frac{1}{\varepsilon}|v^\varepsilon - v|_0 \le C,$$

where v solves (2.1) *and the constant C only depends on K, λ from (A1).*

We are now in a position to state and prove the main result of this section.

Theorem 2.3. *Assume (A1) and (A2). If u and v are the solutions of* (2.2) *and* (2.1) *respectively, then for k small enough,*

$$0 \le v_i - u \le Ck^{1/3} \quad \textit{in} \quad \mathbb{R}^N, \quad i \in \mathcal{I},$$

where C only depends on λ, K from (A1).

Remark 2.1. This seems to be the first time the rate of convergence is obtained for switching system approximations of second-order equations.

Proof of Theorem 2.3. Since $w = (u, \dots, u)$ is a subsolution of (2.1), comparison for (2.1) (Proposition 2.1 (i)) yields $u \le v_i$ for $i \in \mathcal{I}$.

To get the other bound, we use an argument suggested by P.-L. Lions [33] together with the regularization procedure of Krylov [27]. Consider first system (2.3). It follows that, for every $|e| \le \varepsilon$,

$$\sup_{\alpha\in\mathcal{A}_i} \mathcal{L}^\alpha(x+e, v_i^\varepsilon(x), Dv_i^\varepsilon, D^2 v_i^\varepsilon) \le 0 \quad \text{in} \quad \mathbb{R}^N, \quad i \in \mathcal{I}.$$

After a change of variables, we see that for every $|e| \le \varepsilon$, $v^\varepsilon(x-e)$ is a subsolution of the following system of uncoupled equations

$$\sup_{\alpha\in\mathcal{A}_i} \mathcal{L}^\alpha(x, w_i, Dw_i, D^2 w_i) = 0 \quad \text{in} \quad \mathbb{R}^N, \quad i \in \mathcal{I}. \tag{2.4}$$

Define $v_\varepsilon := v^\varepsilon * \rho_\varepsilon$ where $\{\rho_\varepsilon\}_\varepsilon$ is the sequence of mollifiers defined at the end of the introduction. Then v_ε is also a subsolution of (2.4) since it can be viewed as the limit of convex combinations of subsolutions $v^\varepsilon(x-e)$ of the convex system of equations (2.4). We refer to the Appendix in [2] for the details.

On the other hand, since v^ε is a continuous subsolution of (2.3), we have

$$v_i^\varepsilon \le \min_{j\ne i} v_j^\varepsilon + k \quad \text{in} \quad \mathbb{R}^N, \quad i \in \mathcal{I}.$$

It follows that $\max_i v_i^\varepsilon(x) - \min_i v_i^\varepsilon(x) \le k$, and hence

$$|v_i^\varepsilon - v_j^\varepsilon|_0 \le k, \quad i, j \in \mathcal{I}.$$

Then, by the definition and properties of v_ε, we have

$$|D^n v_{\varepsilon i} - D^n v_{\varepsilon j}|_0 \le C\frac{k}{\varepsilon^n}, \quad n \in \mathbb{N}, \quad i, j \in \mathcal{I},$$

where C only depends on ρ. Furthermore, from these bounds, we see that for $\varepsilon < 1$,

$$\left| \sup_{\alpha\in\mathcal{A}_i} \mathcal{L}^\alpha[v_{\varepsilon j}] - \sup_{\alpha\in\mathcal{A}_i} \mathcal{L}^\alpha[v_{\varepsilon i}] \right| \le C\frac{k}{\varepsilon^2} \quad \text{in} \quad \mathbb{R}^N, \quad i, j \in \mathcal{I}.$$

Here C only depends on $|\sigma|_0, |b|_0, |c|_0$ and ρ. Since v_ε is a subsolution of (2.4), this means that,

$$\sup_{\alpha\in\overline{\mathcal{A}}} \mathcal{L}^\alpha(x, v_{\varepsilon i}, Dv_{\varepsilon i}, D^2 v_{\varepsilon i}) \le C\frac{k}{\varepsilon^2} \quad \text{in} \quad \mathbb{R}^N, \quad i \in \mathcal{I}.$$

So by (A1) and the definition of $\mathcal{L}$, we see that $v_{\varepsilon i} - \frac{1}{\lambda} C \frac{k}{\varepsilon^2}$ is a subsolution of equation (2.2).

Comparison for (2.2) (Proposition 2.1 (i)) yields

$$v_{\varepsilon i} - u \le \frac{1}{\lambda} C \frac{k}{\varepsilon^2} \quad \text{in} \quad \mathbb{R}^N, \quad i \in \mathcal{I}.$$

Hence, by properties of mollifiers and the regularity of v_i^ε, we have

$$v_i - u \le v_i - v_{\varepsilon i} + v_{\varepsilon i} - u \le C\varepsilon + \frac{1}{\lambda} C \frac{k}{\varepsilon^2} \quad \text{in} \quad \mathbb{R}^N, \quad i \in \mathcal{I}.$$

Minimizing w.r.t. ε yields the result. □

3. Convergence rate for the HJB equation.

In this section we derive an error bound for the convergence of the solution of the scheme (1.2) to the solution of the HJB equation (1.1). This result is general and derived using only PDE methods, and it extends and improves earlier results by Krylov [26, 27], Barles and Jakobsen [2, 21].

We assume that assumptions (A1) and (A2) of Section 2 hold. As a special case of Proposition 2.1, we have the following well-posedness and regularity result for (1.1):

Proposition 3.1. *(i) Assume (A1). If w_1 and w_2 are sub- and supersolutions of* (1.1)*, then $w_1 \leq w_2$.*

(ii) Assume (A1) and (A2). Then there exists a unique solution u of (1.1) *satisfying*

$$|u|_1 \leq C,$$

where the constant C only depends on K, λ from (A1).

For the scheme (1.2) we assume:

(S1) (Monotonicity) For every $h > 0$, $x \in \mathbb{R}^N$, $r \in \mathbb{R}$, $m \geq 0$, and bounded continuous functions u, v such that $u \leq v$ in $\mathbb{R}^N$, the following holds:

$$S(h, x, r + m, [u + m]_x) \geq \lambda m + S(h, x, r, [v]_x).$$

(S2) (Regularity) For every $h > 0$ and $\phi \in C_b(\mathbb{R}^N)$, $x \mapsto S(h, x, \phi(x), [\phi]_x)$ is bounded and continuous in $\mathbb{R}^N$ and the function $r \mapsto S(h, x, r, [\phi]_x)$ is uniformly continuous for bounded r, uniformly in $x \in \mathbb{R}^N$.

(S3) (Consistency) There exists integers $n, k_i \geq 0$, $i = 1, 2, \ldots, n$, and a constant K_c such that for every $h \geq 0$, $x \in \mathbb{R}^N$, and smooth function ϕ:

$$\left|F(x, \phi(x), D\phi(x), D^2\phi(x)) - S(h, x, \phi(x), [\phi]_x)\right| \leq K_c \sum_{k_i \neq 0} |D^i \phi|_0 \, h^{k_i}.$$

Remark 3.1. Condition (S1) and (S2) imply a comparison result for bounded continuous solutions of (1.2), see [2].

Before we continue, we mention that the upper bound on the error $u - u_h$ is known from [2], see also [27, 21]. Let us state the result here.

Proposition 3.2. *Assume (A1), (A2), (S1) – (S3), and that* (1.2) *has a unique solution $u_h \in C_b(\mathbb{R}^N)$. If u is the solution of* (1.1)*, then, for sufficiently small $h > 0$, we have*

$$u - u_h \leq C h^{\gamma} \quad in \quad \mathbb{R}^N,$$

where $\gamma := \min_{k_i \neq 0} \left\{ \frac{k_i}{i} \right\}$ and C only depends on λ, K, K_c from (A1), (S3).

Remark 3.2. Existence of $u_h \in C_b(\mathbb{R}^N)$ must be proved for each particular scheme S. We refer to [26, 27, 2, 21] for examples of such arguments.

As mentioned in the Introduction, the proof of this proposition relies on the regularization procedure of Krylov which was also used in Section 2. The idea is to obtain a smooth subsolution of equation (1.1) which is close to the solution of this equation. This then yields the upper bound after classical computations. This approach however does not yield the lower bound unless you require much stronger assumptions on the scheme (1.2), see [2, 21, 26].

To avoid such restrictive assumptions, we use a different technique here. The key point is to obtain approximate "almost smooth" supersolutions by considering the following switching system approximation of (1.1):

$$F_i^\varepsilon(x, v^\varepsilon, Dv_i^\varepsilon, D^2 v_i^\varepsilon) = 0 \quad \text{in} \quad \mathbb{R}^N, \quad i \in \mathcal{I}, \tag{3.1}$$

where $v^\varepsilon = (v_1^\varepsilon, \cdots, v_M^\varepsilon)$,

$$F_i^\varepsilon(x, r, p, X) = \max\Big\{ \min_{|e|\le\varepsilon} \mathcal{L}^{\alpha_i}(x, r_i, p, X);\ r_i - \mathcal{M}_i r \Big\},$$

and $\mathcal{L}$ and $\mathcal{M}$ are defined below (1.1) and (2.1) respectively. The solution of this system is expected to be close to the solution of (1.1) if k and ε are small and $\{\alpha_i\}_{i\in\mathcal{I}} \subset \mathcal{A}$ is a sufficiently refined grid for $\mathcal{A}$. In fact for this to be true we need to assume that the coefficients $\sigma^\alpha, b^\alpha, c^\alpha, f^\alpha$ can be approximated uniformly in x by $\sigma^{\alpha_i}, b^{\alpha_i}, c^{\alpha_i}, f^{\alpha_i}$. The precise assumption is:

(A3) For every $\delta > 0$, there are $M \in \mathbb{N}$ and $\{\alpha_i\}_{i=1}^M \subset \mathcal{A}$, such that for any $\alpha \in \mathcal{A}$,

$$\inf_{1\le i\le M} \left(|\sigma^\alpha - \sigma^{\alpha_i}|_0 + |b^\alpha - b^{\alpha_i}|_0 + |c^\alpha - c^{\alpha_i}|_0 + |f^\alpha - f^{\alpha_i}|_0\right) < \delta.$$

Remark 3.3. The typical cases where (A3) is satisfied are (i) when $\mathcal{A}$ is a finite set and (ii) when all coefficients are uniformly continuous in α, uniformly in x.

For equation (3.1), we have the following result.

Lemma 3.3. *Assume (A1) and (A2).*

(a) There exists a unique solution v^ε of (3.1) *satisfying $|v^\varepsilon|_1 \le C$, where C only depends on λ, K from (A1).*

(b) Assume in addition (A3), and let u denote the solution of (1.1)*. Then for any $\delta > 0$ there are $M \in \mathbb{N}$ and $\{\alpha_i\}_{i=1}^M \subset \mathcal{A}$ such that the solution v_ε of* (3.1) *satisfy*

$$\max_i |u - v_i^\varepsilon|_0 \le C(\varepsilon + k^{1/3} + \delta),$$

where C only depends on λ, K from (A1).

The (almost) smooth supersolutions of (1.1) we are looking for are built out of the v_i^ε's by mollification. Before giving the next lemma, we remind the reader that the sequence of mollifiers $\{\rho_\varepsilon\}_\varepsilon$ is defined at the end of the introduction.

Lemma 3.4. *Assume (A1), (A2), and define* $v_{\varepsilon i} := \rho_\varepsilon * v_i^\varepsilon$ *for* $i \in \mathcal{I}$.

(a) There is a constant C *depending only on* λ, K *from (A1), such that*

$$|v_{\varepsilon j} - v_i^\varepsilon|_0 \le C(k+\varepsilon) \quad \text{for} \quad i,j \in \mathcal{I}.$$

(b) Assume in addition that $\varepsilon \le (4\sup_i [v_i^\varepsilon]_1)^{-1}k$. *For every* $x \in \mathbb{R}^N$, *if* $j := \mathrm{argmin}_{i\in\mathcal{I}} v_{\varepsilon i}(x)$, *then*

$$\mathcal{L}^{\alpha_j}(x, v_{\varepsilon j}(x), Dv_{\varepsilon j}(x), D^2 v_{\varepsilon j}(x)) \ge 0 .$$

Lemma 3.4 (b) implies that $w := \min_{i\in\mathcal{I}} v_{\varepsilon i}$ is a viscosity supersolution of (1.1) in all of $\mathbb{R}^N$ (at least this follows from the proof). This function is an "almost smooth" supersolution in the sense that, at any point x, it is only the smooth function $v_{\varepsilon j}$ of Lemma 3.4 (b) (which is a supersolution at this point) which is really playing a role. This can be seen from the proof of the rate of convergence below.

We prove these two lemmas after having stated and proved the main result of this paper – the result giving the lower bound on the error $u - u_h$ for the scheme (1.2).

Theorem 3.5. *Assume (A1) – (A3), (S1), (S3), and that* (1.2) *has a unique solution* $u_h \in C_b(\mathbb{R}^N)$. *If* u *is the solution of* (1.1), *then, for sufficiently small* $h > 0$, *we have*

$$-Ch^{\bar\gamma} \le u - u_h \quad \text{in} \quad \mathbb{R}^N,$$

where $\bar\gamma := \min_{k_i \neq 0} \left\{ \frac{k_i}{3i-2} \right\}$ *and* C *only depends on* λ, K, K_c *from (A1), (S3).*

Proof. We fix a $\delta > 0$ and pick the corresponding $\{\alpha_i\}_{\mathcal{I}}$ according to (A3). Then we consider the solution v^ε of (3.1) corresponding to this choice of $\{\alpha_i\}_{\mathcal{I}}$. Lemma 3.3 yields existence and properties of v^ε. Furthermore, we mollify this function to obtain v_ε as in Lemma 3.4.

We proceed to obtain an estimate for

$$m := \sup_{y\in\mathbb{R}^N} \{u_h(y) - w(y)\},$$

where $w := \min_{i\in\mathcal{I}} v_{\varepsilon i}$. In order to have a "max" instead of a "sup", we approximate m by

$$m_\kappa := \sup_{y\in\mathbb{R}^N} \{u_h(y) - w(y) - \kappa\phi(y)\}, \tag{3.2}$$

where $\kappa > 0$ is a small constant and $\phi(y) = (1+|y|^2)^{1/2}$. Since u_h and w are continuous, it is clear that the supremum (3.2) is attained at some point $x \in \mathbb{R}^N$. Because of the definition of w, it is easy to see that x is also a maximum point of

$$\sup_{y\in\mathbb{R}^N} \{u_h(y) - v_{\varepsilon i}(y) - \kappa\phi(y)\} \tag{3.3}$$

when $i = \mathrm{argmin}_{j\in\mathcal{I}} v_{\varepsilon j}(x)$. Notice that this supremum is still m_κ.

Now take $\varepsilon = (4\sup_i [v_i^\varepsilon]_1)^{-1}k$. From Lemma 3.4 (b), the properties of ϕ, and (A1), we see that

$$\sup_{\alpha\in\mathcal{A}} \mathcal{L}^\alpha(x, (v_{\varepsilon i} + \kappa\phi)(x), D(v_{\varepsilon i} + \kappa\phi)(x), D^2(v_{\varepsilon i} + \kappa\phi)(x)) \geq -C\kappa, \tag{3.4}$$

where C only depends on K from (A1) ($C = \sup_{\alpha,x}\{|\sigma^\alpha|_0^2 + |b^\alpha|\}$).

Let us estimate m_κ. By (3.4) and (S3) we have

$$-C\kappa \leq S(h, x, (v_{\varepsilon i} + \kappa\phi)(x), [v_{\varepsilon i} + \kappa\phi]_x) + K_c \sum_{k_i \neq 0} |D^i(v_{\varepsilon i} + \kappa\phi)|_0 h^{k_i}.$$

By the definitions of $v_{\varepsilon i}$ and ϕ, we can conclude that

$$-C \sum_{k_i\neq 0} \varepsilon^{1-i} h^{k_i} + \mathcal{O}(\kappa) \leq S(h, x, (v_{\varepsilon i} + \kappa\phi)(x), [v_{\varepsilon i} + \kappa\phi]_x), \tag{3.5}$$

where C only depends on ρ and λ, K from (A1). On the other hand, using (S1), (3.3), and the definition of m_κ, we see that

$$\begin{aligned} S(h, x, (v_{\varepsilon i} + \kappa\phi)(x), [v_{\varepsilon i} + \kappa\phi]_x) &\leq S(h, x, u_h(x) - m_\kappa, [u_h - m_\kappa]_x) \\ &\leq -\lambda m_\kappa + S(h, x, u_h(x), [u_h]_x) = -\lambda m_\kappa, \end{aligned}$$

where the last equality follows since u_h is the solution of (1.2). From this inequality and (3.5), we have

$$\lambda m_\kappa \leq C \sum_{k_i\neq 0} \varepsilon^{1-i} h^{k_i} + \mathcal{O}(\kappa).$$

From this estimate, we obtain the estimate for m by sending $\kappa \to 0$ and noting that $m_\kappa \to m$.

Using the estimate for m, we now derive the lower bound on the error. Fix an arbitrary $y \in \mathbb{R}^N$. From the definition of m, we see that

$$\begin{aligned} u_h(y) - u(y) &\leq u_h(y) - v_{\varepsilon i}(y) + v_{\varepsilon i}(y) - u(y) \\ &\leq m + v_{\varepsilon i}(y) - u(y). \end{aligned}$$

Using the bound on m, Lemmas 3.4 (a) and 3.3 (b), we obtain

$$u_h(y) - u(y) \leq C\left(\sum_{k_i\neq 0} \varepsilon^{1-i} h^{k_i} + \varepsilon + k + k^{1/3} + \delta\right).$$

The constant C does not depend on y and therefore the right-hand side is a uniform in y upper bound for $u_h - u$.

The conclusion follows by choosing

$$\varepsilon = \max_{k_i\neq 0} h^{\frac{3k_i}{3i-2}} \quad \text{and} \quad k = 4\sup_i [v_i^\varepsilon]_1 \varepsilon,$$

and sending $\delta \to 0$ (since all constants are independent of the size of $\mathcal{I}$). $\square$

Now we give the proofs of Lemmas 3.3 and 3.4.

Proof of Lemma 3.3.
1. First we approximate (1.1) by the following equation:

$$\sup_{i\in\mathcal{I}} \mathcal{L}^{\alpha_i}(x,v,Dv,D^2v)=0 \quad \text{in} \quad \mathbb{R}^N.$$

From assumption (A3) and Lemmas A.1 and A.3 in the Appendix, we have the following result: There exist a unique solution v of the above equation satisfying

$$|v-u|_0\le C\delta,$$

where C only depends on λ, K from (A1).
2. We continue by approximating the above equation by the following switching system:

$$\max\Big\{\mathcal{L}^{\alpha_i}(x,v_i,Dv_i,D^2v_i);\ v_i-\mathcal{M}_iv\Big\}=0 \quad \text{in} \quad \mathbb{R}^N, \quad i\in\mathcal{I},$$

where $\mathcal{M}$ is defined below (2.1). From Proposition 2.1 and Theorem 2.3 in Section 2 we have existence and uniqueness of a solution $\bar v$ of the above system satisfying

$$|\bar v_i - v|_0\le Ck^{1/3}, \quad i\in\mathcal{I},$$

where C only depends on ρ and λ, K from (A1).
3. The switching system defined in the previous step is nothing but (3.1) with $\varepsilon=0$ or (2.3) with the $\mathcal{A}_i$'s being singletons. Proposition 2.2 in Section 2 yields the existence and uniqueness of a solution v^ε of (3.1) satisfying

$$|v^\varepsilon|_1+\frac{1}{\varepsilon}|v^\varepsilon-\bar v|_0\le C,$$

where C only depends on λ, K from (A1).
4. The proof is complete by combining the estimates in steps 1 – 3, and noting that (A3) is only needed in step 1. □

Proof of Lemma 3.4. We start by (a). From the properties of mollifiers and the Lipschitz continuity of v^ε, it is immediate that

$$|v_{\varepsilon i}-v_i^\varepsilon|_0\le C\varepsilon, \quad i\in\mathcal{I},$$

where $C=\max_i[v_i^\varepsilon]_1$ depends only on K,λ from (A1). Furthermore we saw in the proof of Theorem 2.3 in Section 2 that

$$0\le \max_i v_i^\varepsilon-\min_i v_i^\varepsilon\le k \quad \text{in} \quad \mathbb{R}^N.$$

From these two estimates, (a) follows.

Now consider (b). We consider an arbitrary point $x\in\mathbb{R}^N$ and set

$$j=\mathrm{argmin}_{i\in\mathcal{I}}v_{\varepsilon i}(x).$$

Then, by definition of $\mathcal{M}$ and j, we have

$$v_{\varepsilon j}(x)-\mathcal{M}_jv_\varepsilon(x)=\max_{i\ne j}\big\{v_{\varepsilon j}(x)-v_{\varepsilon i}(x)-k\big\}\le -k.$$

Part (a) then leads to

$$v_j^\varepsilon(x)-\mathcal{M}_jv^\varepsilon(x)\le -k+2[v_j^\varepsilon]_1\varepsilon,$$

and by using the Lipschitz continuity of v^ε (Lemma 3.3),

$$v_j^\varepsilon(y) - \mathcal{M}_j v^\varepsilon(y) \leq -k + 2[v_j^\varepsilon]_1(\varepsilon + |x-y|).$$

From this we conclude that if $|x-y| < \varepsilon$ and $\varepsilon \leq (4\max_i [v_i^\varepsilon]_1)^{-1}k$, then

$$v_j^\varepsilon(y) - \mathcal{M}_j v^\varepsilon(y) < 0.$$

Equation (3.1) then implies

$$\inf_{|e|\leq\varepsilon} \mathcal{L}^{\alpha_j}(y+e, v_j^\varepsilon(y), Dv_j^\varepsilon(y), D^2 v_j^\varepsilon(y)) = 0.$$

After a change of variables we see that for every $|e| \leq \varepsilon$,

$$\mathcal{L}^{\alpha_j}(x, v_j^\varepsilon(x-e), Dv_j^\varepsilon(x-e), D^2 v_j^\varepsilon(x-e)) \geq 0. \tag{3.6}$$

In other words, for every $|e| \leq \varepsilon$, $v_j^\varepsilon(x-e)$ is a supersolution at x of

$$\mathcal{L}^{\alpha_j}(x, w, Dw, D^2 w) = 0. \tag{3.7}$$

By mollifying (3.6) we see formally that $v_{\varepsilon j}$ is also a supersolution of (3.7) at x and hence a (viscosity) supersolution of the HJB equation (1.1) at x. This is correct since $v_{\varepsilon j}$ can be viewed as the limit of convex combinations of supersolutions $v_j^\varepsilon(x-e)$ of the linear and hence concave equation (3.7), we refer to the Appendix in [2] for the details. We conclude the proof by noting that since $v_{\varepsilon j}$ is smooth, it is in fact a classical supersolution of (1.1) at x. □

4. Monotone Finite Difference Methods.

As an application of the results in the previous section we derive here a the rate of convergence for a finite difference scheme proposed by Kushner [11, 14] for the N-dimensional HJB equation (1.1). The notation for these schemes is taken from [11, 14]. We start by naming the difference operators we need. Let $\{e_i\}_{i=1}^N$ denote the standard basis in $\mathbb{R}^N$ and define

$$\begin{aligned}
\Delta_{x_i}^{\pm} w(x) &= \pm\frac{1}{h}\{w(x \pm e_i h) - w(x)\},\\
\Delta_{x_i}^2 w(x) &= \frac{1}{h^2}\{w(x+e_i h) - 2w(x) + w(x-e_i h)\},\\
\Delta_{x_i x_j}^{+} w(x) &= \frac{1}{2h^2}\{2w(x) + w(x+e_i h+e_j h) + w(x-e_i h-e_j h)\}\\
&\quad - \frac{1}{2h^2}\{w(x+e_i h) + w(x-e_i h) + w(x+e_j h) + w(x-e_j h)\},\\
\Delta_{x_i x_j}^{-} w(x) &= \frac{1}{2h^2}\{w(x+e_i h) + w(x-e_i h) + w(x+e_j h) + w(x-e_j h)\}\\
&\quad - \frac{1}{2h^2}\{2w(x) + w(x+e_i h-e_j h) + w(x-e_i h+e_j h)\}.
\end{aligned}$$

Now we define the schemes as follows,

$$\tilde{F}(x, u_h(x), \Delta_{x_i}^{\pm} u_h(x), \Delta_{x_i}^2 u_h(x), \Delta_{x_i x_j}^{\pm} u_h(x)) = 0, \tag{4.1}$$

where

$$\tilde{F}(x,t,p_i^\pm,A_{ii},A_{ij}^\pm) = \sup_{\alpha\in\mathcal{A}} \Big\{ \sum_{i=1}^N \Big[-\frac{a_{ii}^\alpha}{2}A_{ii} + \sum_{j\neq i}\Big(-\frac{a_{ij}^{\alpha+}}{2}A_{ij}^+ + \frac{a_{ij}^{\alpha-}}{2}A_{ij}^-\Big) \\ - b_i^{\alpha+}(x)p_i^+ + b_i^{\alpha-}(x)p_i^-\Big] + c^\alpha(x)t - f^\alpha(x)\Big\},$$

and $b^+ = \max\{b,0\}$ and $b^- = (-b)^+$ $(b = b^+ - b^-)$.

Assume that (A1) holds. In order to obtain the required monotonicity of these schemes, we need to assume in addition that the matrix a is diagonally dominant,

$$a_{ii}^\alpha(x) - \sum_{j\neq i}|a_{ij}^\alpha(x)| \geq 0 \quad \text{in} \quad \mathbb{R}^N, \quad i=1,\dots,N. \tag{4.2}$$

We also assume that the coefficients are normalized so that

$$\sum_{i=1}^N \Big\{ a_{ii}^\alpha(x) - \sum_{j\neq i}|a_{ij}^\alpha(x)| + |b_i^\alpha(x)|\Big\} \leq 1 \quad \text{in} \quad \mathbb{R}^N. \tag{4.3}$$

Assumption (4.2) is standard in numerical analysis, see [11, 14]. We also refer to Lions and Mercier [34] and to Bonnans and Zidani [4] for a discussion of this condition. Assumption (4.3) is always satisfied after a multiplication in (1.1) by an appropriate positive constant.

From the results in Section 3, we have the following bound on $u - u_h$:

Theorem 4.1. *Assume (A1) – (A3),* (4.2), *and* (4.3) *hold. If u and $u_h \in C_b(\mathbb{R}^N)$ are solutions of* (1.1) *and* (4.1) *respectively, then for $h > 0$ sufficiently small,*

$$|u - u_h|_0 \leq Ch^{1/5}.$$

Remark 4.1. Krylov [27] obtains the rate 1/27 using probabilistic methods. One contribution of this paper is to improve this rate to 1/5.

By Theorems 3.2 and 3.5 in Section 3, the above result holds if we can define S in (1.2), check that assumptions (S1) – (S3) hold with $k_2 = 1$, $k_4 = 2$, and $k_i = 0$ otherwise, and prove existence of $u_h \in C_b(\mathbb{R}^N)$. Let us proceed to write down S. In order to better see the monotonicity of the scheme and to fix some more notation, we are going to rewrite (4.1) as a discrete dynamical programming principle. We refer to [11] for the probabilistic interpretation. Define the following one step transition probabilities,

$$p^\alpha(x,x) = 1 - \sum_{i=1}^N \Big\{ a_{ii}^\alpha(x) - \sum_{j\neq i}|a_{ij}^\alpha(x)| + h|b_i^\alpha(x)|\Big\},$$

$$p^\alpha(x,x\pm e_ih) = \frac{a_{ii}^\alpha(x)}{2} - \sum_{j\neq i}\frac{|a_{ij}^\alpha(x)|}{2} + hb_i^{\alpha\pm}(x),$$

$$p^\alpha(x, x+e_ih \pm e_jh) = \frac{a_{ij}^{\alpha\pm}(x)}{2},$$
$$p^\alpha(x, x-e_ih \pm e_jh) = \frac{a_{ij}^{\alpha\mp}(x)}{2},$$

and $p^\alpha(x,y) = 0$ for all other y. Note that by (4.2) and (4.3), $0 \le p^\alpha(x,y) \le 1$ for all α, x, y if $h \le 1$. Furthermore $\sum_{z\in h\mathbb{Z}^N} p^\alpha(x, x+z) = 1$ for all α, x. Tedious but straightforward computations show that the following equation is equivalent to (4.1),

$$u_h(x) = \inf_{\alpha\in\mathcal{A}} \left\{ \frac{1}{1+h^2c^\alpha(x)} \Big(\sum_{z\in h\mathbb{Z}^N} p^\alpha(x,x+z)u_h(x+z) + h^2 f^\alpha(x) \Big) \right\}.$$

This is the discrete dynamical programming principle. From this equation we define S. For $\phi \in C_b(\mathbb{R}^N)$, set $[\phi]_x^h(\cdot) := \phi(x+\cdot)$ and

$$S(h, y, t, [\phi]_x^h) := \sup_{\alpha\in\mathcal{A}} \left\{ -\frac{1}{h^2} \left[\sum_{z\in h\mathbb{Z}^N} p^\alpha(y, y+z)[\phi]_x^h(z) - t \right] + c^\alpha(x)t - f^\alpha(y) \right\}.$$

Using this definition of S, it is easy to check (S1) – (S3), see the lemma below (see also [2]). Existence of solutions $u_h \in C_b(\mathbb{R}^N)$ of (4.1) can be proved using the contraction mapping theorem, we refer to [26, 27, 2, 21] for such arguments. Thus, we may conclude that Theorem 4.1 holds.

Lemma 4.2. *Assume (A1), (A2),* (4.2), (4.3)*, and* $0 < h < 1$*. Then the scheme* (4.1) *satisfy conditions (S1) – (S3), where (S3) takes the form*

$$|F(x, v, Dv, D^2v) - S(h, x, v(x), [v]_x)| \le \sup_\alpha |b^\alpha|_0 |D^2v|_0 h + \sup_\alpha |\sigma^\alpha|_0^2 |D^4v|_0 h^2.$$

5. Extensions and Remarks

Let us first consider the case when **(A2)** is not satisfied. Then the solutions of the different equations are only Hölder continuous. E.g. for the HJB equation (1.1) we have the following result:

Lemma 5.1. *Assume (A1) and define* $\lambda_0 := \sup_{\mathcal{A}}\{[\sigma]_1^2 + [b]_1\}$*. If* $\lambda < \lambda_0$*, then there exist unique solution* $u \in C^{0,\delta}(\mathbb{R}^N)$ *of* (1.1)*, where* $\delta = \lambda/\lambda_0$*.*

This result was proved in [30]. We claim that under (A1), we have the same regularity (the same δ) for all equations considered in this paper. We skip the tedious proof of this claim. In the rest of this section, the solutions of the different equations are assumed to belong to $C^{0,\delta}(\mathbb{R}^N)$ with the same fixed $\delta \in (0,1]$.

Lower than Lipschitz regularity of solutions implies lower convergence rates than obtained in Sections 2 – 4. We will now state the Hölder version of these results without proofs. The proofs are not much different from the proofs given above, and moreover, the Hölder case was extensively studied in [2]. We start by the convergence rate for the switching system approximation of Section 2.

Proposition 5.2. *Assume (A1). If* $\bar{u}$ *and* v *are the solutions of* (2.2) *and* (2.1) *belonging to* $C^{0,\delta}(\mathbb{R}^N)$, *then for* k *small enough,*

$$0 \leq v_i - \bar{u} \leq Ck^{\frac{\delta}{2+\delta}} \quad in \quad \mathbb{R}^N, \quad i \in \mathcal{I},$$

where C *only depends on* λ, K *from (A1).*

The upper bound on the error for monotone approximation schemes (1.2) for the HJB equation (1.1) is given by the following result.

Proposition 5.3. *Assume (A1), (S1) – (S3), and that* (1.2) *has a unique solution* $u_h \in C_b(\mathbb{R}^N)$. *If* $u \in C^{0,\delta}(\mathbb{R}^N)$ *is the solution of* (1.1), *then for sufficiently small* $h > 0$, *we have*

$$u - u_h \leq Ch^{\delta\gamma} \quad in \quad \mathbb{R}^N,$$

where γ *and* C *are defined in Proposition 3.2.*

This proposition was essentially proved in [2], see [21] for this form of the result. Finally, we have come to the Hölder version of the main result of this paper:

Proposition 5.4. *Assume (A1), (A3), (S1), (S3), and that* (1.2) *has a unique solution* $u_h \in C_b(\mathbb{R}^N)$. *If* $u \in C^{0,\delta}(\mathbb{R}^N)$ *is the solution of* (1.1), *then for sufficiently small* $h > 0$, *we have*

$$-Ch^{\bar{\gamma}} \leq u - u_h \quad in \quad \mathbb{R}^N,$$

where $\bar{\gamma} := \min_{k_i \neq 0} \left\{ \frac{\delta^2 k_i}{(2+\delta)i - 2\delta} \right\}$ *and* C *only depends on* λ, K, K_c *from (A1), (S3).*

Remark 5.1. Above we removed assumption (A2). It is also possible to weaken assumption (A1) by assuming that c, f are only Hölder continuous. This would then lead to Hölder continuous solutions with lower Hölder exponents than above. The above results would continue to hold however, but now with a different δ. We refer to [2] for results in this direction.

Next, we comment on a possible extension to the non-convex/non-concave case. We are interested in the Isaacs equations coming from stochastic differential games,

$$F(x, u, Du, D^2u) = 0 \quad \text{in} \quad \mathbb{R}^N, \tag{5.1}$$

where

$$F(x, t, p, X) = \sup_{\alpha \in \mathcal{A}} \inf_{\beta \in \mathcal{B}} \mathcal{L}^{\alpha,\beta}(x, t, p, X),$$
$$\mathcal{L}^{\alpha,\beta}(x, t, p, X) = -\text{tr}[a^{\alpha,\beta}(x)X] - b^{\alpha,\beta}(x)p + c^{\alpha,\beta}(x)t - f^{\alpha,\beta}(x),$$

and $\mathcal{A}, \mathcal{B}$ are compact metric spaces. Assume that assumptions like (A1) – (A3) are satisfied for this problem. In this case we have well-posedness and Lipschitz regularity results for (5.1) (see the Appendix).

Let $\{\alpha_i\}_{i=1}^M \subset \mathcal{A}$ be a suitable refined grid for $\mathcal{A}$, and consider the question of finding the rate of convergence for the following switching system approximation of (5.1):

$$F_i(x, v, Dv_i, D^2 v_i) = 0 \quad \text{in} \quad \mathbb{R}^N, \quad i \in \mathcal{I} := \{1, \dots, M\}, \tag{5.2}$$

where $v = (v_1, \cdots, v_M)$,

$$F_i(x, r, p, M) = \max\Big\{ \inf_{\beta \in \mathcal{B}} \mathcal{L}^{\alpha_i,\beta}(x, r_i, p, X);\ r_i - \mathcal{M}_i r \Big\},$$

and $\mathcal{M}$ is defined just below (2.1) in Section 2. To the best of our knowledge, this question is still an open problem, and clearly the method used in Section 2 cannot be extended to this case.

However, if we assume that this question has been resolved, then the proof of Theorem 3.5 can be extended to give a lower bound for the error of approximation schemes for (5.1). The only problem we face here, is to extend the proof of Proposition 3.4 (b). But this is trivial because of the concavity of the function $\inf_{\beta \in \mathcal{B}} \mathcal{L}^{\alpha_i,\beta}(x, t, p, X)$.

To get the upper bound on the error, we only need to assume that the Isaacs condition is satisfied, i.e.

$$\sup_{\alpha \in \mathcal{A}} \inf_{\beta \in \mathcal{B}} \mathcal{L}^{\alpha,\beta}(x, t, p, X) = \inf_{\beta \in \mathcal{B}} \sup_{\alpha \in \mathcal{A}} \mathcal{L}^{\alpha,\beta}(x, t, p, X),$$

for any $x \in \mathbb{R}^N$, $t \in \mathbb{R}$, $p \in \mathbb{R}^N$ and $X \in \mathcal{S}^N$. The upper bound can then be obtained by a symmetric argument, changing "sup" to "inf", "max" to "min" and conversely.

Thus, the rate of convergence of approximation schemes for Isaacs equations would follow from our method if the rate of convergence of the corresponding switching system can be obtained.

Appendix A. Well-posedness, regularity, and continuous dependence for switching systems.

In this section we give well-posedness, regularity, and continuous dependence results for solutions of a very general switching system that has as special cases the scalar HJB and Isaacs equations (1.1) and (5.1), and the switching systems (2.1), (2.3), (3.1), (5.2).

We consider the following system:

$$F_i(x, u, Du_i, D^2 u_i) = 0 \quad \text{in} \quad \mathbb{R}^N, \quad i \in \mathcal{I} := \{1, \dots, M\}, \tag{A.1}$$

with

$$F_i(x, r, p, X) = \max\Big\{ \sup_{\alpha \in \mathcal{A}} \inf_{\beta \in \mathcal{B}} \mathcal{L}_i^{\alpha,\beta}(x, r_i, p, X);\ r_i - \mathcal{M}_i r \Big\},$$

$$\mathcal{L}_i^{\alpha,\beta}(x, t, p, X) = -\mathrm{tr}[a_i^{\alpha,\beta}(x) X] - b_i^{\alpha,\beta}(x) p + c_i^{\alpha,\beta}(x) t - f_i^{\alpha,\beta}(x),$$

where $\mathcal{M}$ is defined below (2.1), $\mathcal{A}, \mathcal{B}$ are compact metric spaces, r is a vector $r = (r_1, \dots, r_M)$, and $k > 0$ is a constant (the switching cost). See [12, 6, 38, 18, 17] for more information about such systems.

We make the following assumptions:

(A1) For any α,β,i, $a_i^{\alpha,\beta}=\frac{1}{2}\sigma_i^{\alpha,\beta}\sigma_i^{\alpha,\beta T}$ for some $N\times P$ matrix $\sigma_i^{\alpha,\beta}$. Furthermore, there are constants λ, C independent of i,α,β, such that

$$c\geq\lambda>0\quad\text{and}\quad[\sigma_i^{\alpha,\beta}]_1+[b_i^{\alpha,\beta}]_1+[c_i^{\alpha,\beta}]_1+|f_i^{\alpha,\beta}|_1\leq C.$$

(A2) The constant λ in (A1) satisfy $\lambda>\sup_{i,\alpha,\beta}\Big\{[\sigma_i^{\alpha,\beta}]_1^2+[b_i^{\alpha,\beta}]_1\Big\}$.

We start by comparison, existence, uniqueness, and L^∞ bounds on the solution and its gradient. Before stating the results, we define $USC(\mathbb{R}^N;\mathbb{R}^M)$ and $LSC(\mathbb{R}^N;\mathbb{R}^M)$ to be the spaces of upper and lower semi-continuous functions from $\mathbb{R}^N$ into $\mathbb{R}^M$ respectively.

Theorem A.1. *Assume (A1) holds.*

(i) If $u\in USC(\mathbb{R}^N;\mathbb{R}^M)$ is a subsolution of (A.1) *bounded above and $v\in LSC(\mathbb{R}^N;\mathbb{R}^M)$ supersolution of* (A.1) *bounded below, then $u\leq v$ in $\mathbb{R}^N$.*

(ii) There exists a unique bounded continuous solution u of (A.1) *satisfying*

$$\max_i |u_i|_0\leq\sup_{i,\alpha,\beta}\frac{|f_i^{\alpha,\beta}|_0}{\lambda},$$

(iii) If in addition (A2) holds, then u is Lipschitz continuous and

$$\max_i\,[u_i]_1\leq\sup_{i,\alpha,\beta}\frac{|u^i|_0[c_i^{\alpha,\beta}]_1+[f_i^{\alpha,\beta}]_1}{\lambda-[\sigma_i^{\alpha,\beta}]_1^2-[b_i^{\alpha,\beta}]_1}.$$

Remark A.1. These bounds have the same form as for linear equations [15] and HJB equations [30].

Before giving the proof we state and prove a key technical lemma.

Lemma A.2. *Let $u\in USC(\mathbb{R}^N;\mathbb{R}^M)$ be a bounded above subsolution of* (A.1) *and $\bar{u}\in LSC(\mathbb{R}^N;\mathbb{R}^M)$ be a bounded below supersolution of an other equation* (A.1) *where the functions $\mathcal{L}_i^{\alpha,\beta}$ are replaced by functions $\bar{\mathcal{L}}_i^{\alpha,\beta}$ satisfying the same assumptions. Let $\phi\in C^2(\mathbb{R}^{2N})$ be a function bounded from below. We denote by*

$$\psi_i(x,y)=u_i(x)-\bar{u}_i(y)-\phi(x,y)\ ,$$

and $M=\sup_{i,x,y}\psi_i(x,y)$. If there exists a maximum point for M, i.e. a point (i',x_0,y_0) such that $\psi_{i'}(x_0,y_0)=M$, then there exists $i_0\in\mathcal{I}$ such that (i_0,x_0,y_0) is also a maximum point for M, and, in addition $\bar{u}_{i_0}(y_0)<\mathcal{M}_{i_0}\bar{u}(y_0)$.

Loosely speaking this lemma means that whenever we do doubling of variables for systems of the type (A.1), we can ignore the $u_i-\mathcal{M}_iu$ part of the equations. So we are more or less back in the scalar case with equations $\sup_\alpha\inf_\beta\mathcal{L}_{i_0}^{\alpha,\beta}[u^{i_0}]\leq 0$ and $\sup_\alpha\inf_\beta\bar{\mathcal{L}}_{i_0}^{\alpha,\beta}[\bar{u}^{i_0}]\geq 0$.

Proof of Lemma A.2. The proof is a "no-loop" argument taken from Ishii and Koike [18]. We assume by contradiction that $\bar{u}_j(y_0) \geq \mathcal{M}_j\bar{u}(y_0)$ for every $j \in A$, where A is the set of j's such that (j, x_0, y_0) is a maximum point for ψ.

We pick a $j \in A$. By the definition of $\mathcal{M}_j$, there is $l \in \mathcal{I}$ such that

$$\mathcal{M}_j\bar{u}(y_0) = \bar{u}_l(y_0) + k\ .$$

By assumption, we have $\bar{u}_j(y_0) \geq \bar{u}_l(y_0) + k$. On the other hand, since u is a subsolution of (A.1), it follows that

$$u_j(x_0) \leq \mathcal{M}_j u(x_0) \leq u_l(x_0) + k.$$

Combining these inequalities yields

$$u_j(x_0) - \bar{u}_j(y_0) \leq k \leq u_l(x_0) - \bar{u}_l(y_0).$$

These inequalities first implies that $l \in A$ and therefore the last inequality is an equality. This, again, implies $\bar{u}_j(y_0) = \bar{u}_l(y_0) + k$.

Since A is finite we may find $j_1, \dots, j_K \in A$ such that $\bar{u}_{j_i}(y_0) = \bar{u}_{j_{i+1}}(y_0) + k$ for $i = 1, \dots, K-1$ and (importantly!) $j_1 = j_K$. But now

$$0 = \sum_{i=1}^{K-1} \big(\bar{u}_{j_i}(y_0) - \bar{u}_{j_{i+1}}(y_0)\big) = (K-1)k > 0,$$

which is a contradiction. The proof is complete. □

Proof of Theorem A.1. Comparison, uniqueness, and existence is proved in [18] for the Dirichlet problem for (1.1) on a bounded domain under assumptions that are satisfied for our problem. The key point here is the comparison principle. To extend this result to an unbounded domain, we only need to modify the test function used in [18] in the standard way. The proof remains practically unchanged.

Let

$$M := \sup_{i,\alpha,\beta} \frac{|f_i^{\alpha,\beta}|_0}{\lambda},$$

then the bound on $|u|_0$ follows from the comparison principle after checking that M $(-M)$ is a supersolution (subsolution) of (A.1). To get the bound on the gradient of u, consider

$$m := \sup_{i,x,y \in \mathbb{R}^N} \left\{ u_i(x) - u_i(y) - L|x-y| \right\}.$$

If by setting

$$L := \sup_{i,\alpha,\beta} \frac{|u_i|_0 [c_i^{\alpha,\beta}]_1 + [f_i^{\alpha,\beta}]_1}{\lambda - [\sigma_i^{\alpha,\beta}]_1^2 - [b_i^{\alpha,\beta}]_1},$$

we can conclude that $m \leq 0$, then we are done. Assume for simplicity that the maximum is attained in $\bar{x}, \bar{y}$. If $\bar{x} = \bar{y}$ then $m = 0$ and we are done. If not, then $L|x-y|$ is smooth at $\bar{x}, \bar{y}$ and a doubling of variables argument leads the $m \leq 0$. This argument is standard after an application of Lemma A.2 which reduce the problem to a scalar problem (see also the proof of

Theorem A.3). We refer the appendix of [15] for the details in the (linear) scalar case. Since the maximum need not be attained, we must modify the test function in the standard way. We skip the details. □

Now we proceed to obtain continuous dependence on the coefficients.

Theorem A.3. *Let u and $\bar{u}$ be solutions of* (A.1) *with coefficients σ, b, c, f and $\bar{\sigma}, \bar{b}, \bar{c}, \bar{f}$ respectively. If both sets of coefficients satisfy (A1) with the same λ, and $|u|_1 + |\bar{u}|_1 \leq M < \infty$, then*

$$\lambda \max_i |u_i - \bar{u}_i|_0 \leq K \sup_{i,\alpha,\beta} |\sigma - \bar{\sigma}|_0 + \sup_{i,\alpha,\beta} \Big\{ 2M|b - \bar{b}|_0 + M|c - \bar{c}|_0 + |f - \bar{f}|_0 \Big\},$$

where

$$K^2 \leq 8M \sup_{i,\alpha,\beta} \Big\{ 2M[\sigma]_1^2 \wedge [\bar{\sigma}]_1^2 + 2M[b]_1 \wedge [\bar{b}]_1 + M[c]_1 \vee [\bar{c}]_1 + [f]_1 \wedge [\bar{f}]_1 \Big\}.$$

Outline of proof. Define

$$m := \sup_{i,x,y} \psi^i(x,y) := \sup_{i,x,y} \left\{ u_i(x) - \bar{u}_i(y) - \frac{1}{\delta}|x-y|^2 - \varepsilon(|x|^2 + |y|^2) \right\}.$$

By the assumptions the supremum is attained at some point (i_0, x_0, y_0). By Lemma A.2, the index i_0 may be chosen so that $\bar{u}_{i_0}(y_0) < \mathcal{M}_{i_0}\bar{u}(y_0)$. With this in mind, the maximum principle for semi continuous functions [8, 9] and the definition of viscosity solutions imply the following inequality:

$$\sup_\alpha \inf_\beta \mathcal{L}_{i_0}^{\alpha,\beta}(x_0, u_{i_0}, p_x, X) - \sup_\alpha \inf_\beta \bar{\mathcal{L}}_{i_0}^{\alpha,\beta}(y_0, \bar{u}_{i_0}, p_y, Y) \leq 0,$$

where $(p_x, X) \in \overline{D}^{2,+} u_{i_0}(x_0)$ and $(p_y, Y) \in \overline{D}^{2,-} \bar{u}_{i_0}(y_0)$ (see [8, 9] for the notation). Furthermore $p_x = \frac{2}{\delta}(x_0 - y_0) + 2\varepsilon x_0$, $p_y = \frac{2}{\delta}(x_0 - y_0) - 2\varepsilon y_0$, and

$$\begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix} \leq \frac{2}{\delta} \begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + 2\varepsilon \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} + \mathcal{O}(\kappa),$$

for some $\kappa > 0$. In the end we will fix δ and ε and send $\kappa \to 0$, so we simply ignore the $\mathcal{O}(\kappa)$-term in the following. The first inequality implies

$$0 \leq \sup_{i,\alpha,\beta} \Big\{ - \operatorname{tr}[\bar{a}(y_0)Y] + \operatorname{tr}[a(x_0)X] + \bar{b}(y_0)p_x - b(x_0)p_y + \bar{c}(y_0)\bar{u}(y_0) - c(x)u(x_0) + \bar{f}(y_0) + f(x_0) \Big\},$$

Note that Lipschitz regularity of the solutions and a standard argument yields

$$|x_0 - y_0| \leq \delta M.$$

So using Ishii's trick on the 2nd order terms [16, pp. 33,34], and a few other manipulations, we get

$$0 \leq \sup_{i,\alpha,\beta} \Big\{ \frac{2}{\delta}|\sigma(x_0) - \bar{\sigma}(y_0)|^2 + 2M|b(x_0) - \bar{b}(y_0)| + C\varepsilon(1 + |x_0|^2 + |y_0|^2)$$

$$+ M|c(x_0) - \bar{c}(y_0)| - \lambda m + |f(x_0) - \bar{f}(y_0)|\Big\}.$$

Some more work lead to an estimate for m depending on δ and ε, and using the definition of m, we obtain an upper bound for $u - \bar{u}$. We finish the proof of the upper bound on $u - \bar{u}$ by minimizing this expression w.r.t. δ and sending $\varepsilon \to 0$. The lower bound follows in a similar fashion. □

Remark A.2. For more details on such manipulations, we refer to [22, 23].

References

[1] R. Abgrall. Numerical discretization of the first-order Hamilton-Jacobi equation on triangular meshes. *Comm. Pure Appl. Math.* 49(12):1339–1373, 1996.

[2] G. Barles and E. R. Jakobsen. On the convergence rate of approximation schemes for Hamilton-Jacobi-Bellman equations. *M2AN Math. Model. Numer. Anal.* 36(1):33–54, 2002.

[3] G. Barles and P. E. Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic Anal.* 4(3):271–283, 1991.

[4] F. Bonnans and H. Zidani. Consistency of generalized finite difference schemes for the stochastic HJB equation. *SIAM J. Numer. Anal.* 41(3), 1008-1021, 2003.

[5] F. Camilli and M. Falcone. An approximation scheme for the optimal control of diffusion processes.

[6] I. Capuzzo-Dolcetta and L. C. Evans. Optimal switching for ordinary differential equations. *SIAM J. Control Optim.* 22(1):143–161, 1984.

[7] I. Capuzzo-Dolcetta and H. Ishii. Approximate solutions of the Bellman equation of deterministic control theory. *Appl. Math. Optim.* 11(2):161–181, 1984.

[8] M. G. Crandall and H. Ishii. The maximum principle for semicontinuous functions. *Differential Integral Equations*, 3(6):1001–1014, 1990.

[9] M. G. Crandall, H. Ishii, and P.-L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)*, 27(1):1–67, 1992.

[10] M. G. Crandall and P.-L. Lions. Two approximations of solutions of Hamilton-Jacobi equations. *Math. Comp.* 43(167):1–19, 1984.

[11] P. Dupuis and H. J. Kushner *Numerical methods for for stochastic control problems in continuous time.* Springer-Verlag, New York, 2001.

[12] L. C. Evans and A. Friedman. Optimal stochastic switching and the Dirichlet problem for the Bellman equation. *Trans. Amer. Math. Soc.* 253:365-389, 1979.

[13] M. Falcone. A numerical approach to the infinite horizon problem of deterministic control theory. *Appl. Math. Optim.* 15(1):1–13, 1987.

[14] W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions.* Springer-Verlag, New York, 1993.

[15] H. Ishii. On the equivalence of two notions of weak solutions, viscosity solutions and distribution solutions. *Funkcial. Ekvac.*, 38(1):101–120, 1995.

[16] H. Ishii. On uniqueness and existence of viscosity solutions of fully nonlinear second-order elliptic PDEs. *Comm. Pure Appl. Math.*, 42(1):15–45, 1989.

[17] H. Ishii and S. Koike. Viscosity solutions for monotone systems of second-order elliptic PDEs. *Comm. Partial Differential Equations* 16(6-7):1095-1128, 1991.

[18] H. Ishii and S. Koike. Viscosity solutions of a system of nonlinear second-order elliptic PDEs arising in switching games. *Funkcial. Ekvac.*, 34:143–155, 1991.

[19] E. R. Jakobsen. Some remarks on penalization and approximation schemes for obstacle problems. In preparation.

[20] E. R. Jakobsen. Error bounds for monotone approximation schemes for non-convex degenerate elliptic equations in $\mathbb{R}^1$. Preprint.

[21] E. R. Jakobsen. On the rate of convergence of approximation schemes for Bellman equations associated with optimal stopping time problems. *Math. Models Methods Appl. Sci. (M3AS)* 13(5):613-644, 2003.

[22] E. R. Jakobsen and K. H. Karlsen. Continuous dependence estimates for viscosity solutions of fully nonlinear degenerate parabolic equations. *J. Differential Equations* 183:497-525, 2002.

[23] E. R. Jakobsen and K. H. Karlsen. Continuous dependence estimates for viscosity solutions of fully nonlinear degenerate elliptic equations. *Electron. J. Diff. Eqns.* 2002(39):1–10, 2002.

[24] E. R. Jakobsen, K. H. Karlsen, and N. H. Risebro. On the convergence rate of operator splitting for Hamilton-Jacobi equations with source terms. *Siam J. Numer. Anal.* Vol. 39(2):499-518, 2001.

[25] G. Kossioris, C. Makridakis, and P. Souganidis. Finite volume schemes for Hamilton-Jacobi equations. *Numer. Math.* 83(3):427–442, 1999.

[26] N. V. Krylov. On the rate of convergence of finite-difference approximations for Bellman's equations. *St. Petersburg Math. J.*, 9(3):639–650, 1997.

[27] N. V. Krylov. On the rate of convergence of finite-difference approximations for Bellman's equations with variable coefficients. *Probab. Theory Ralat. Fields*, 117:1–16, 2000.

[28] O. Lepsky. Spectral viscosity approximations to Hamilton-Jacobi solutions. *SIAM J. Numer. Anal.* 38(5):1439–1453, 2000.

[29] C.-T. Lin and E. Tadmor. L^1-stability and error estimates for approximate Hamilton-Jacobi solutions. *Numer. Math.* 87(4):701–735, 2001.

[30] P.-L. Lions. Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations, Part I: The dynamic programming principle and applications. Comm. P.D.E. **8** (1983).

[31] P.-L. Lions. Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations, Part II: Viscosity solutions and uniqueness. Comm.P.D.E. **8** (1983).

[32] P.-L. Lions. Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations, Part III, Nonlinear PDE and Appl., Séminaire du Collège de France, vol V, Pitman (1985).

[33] P.-L. Lions. Personal communication, 1986.

[34] P.-L. Lions and B. Mercier. Approximation numérique des équations de Hamilton-Jacobi-Bellman. *RAIRO Anal. Numér.* 14(4):369–393, 1980.

[35] P. Soravia. Estimates of convergence of fully discrete schemes for the Isaacs equation of pursuit-evasion differential games via maximum principle. *SIAM J. Control Optim.* 36(1):1–11, 1998.

[36] P. E. Souganidis. Approximation schemes for viscosity solutions of Hamilton-Jacobi equations. *J. Differential Equations* 59(1):1–43, 1985.

[37] P. E. Souganidis. Max-min representations and product formulas for the viscosity solutions of Hamilton-Jacobi equations with applications to differential games. *Nonlinear Anal.* 9(3):217–257, 1985.

[38] N. Yamada. Viscosity solutions for a system of elliptic inequalities with bilateral obstacles. *Fuckcial. Ekvac.* 30(2-3):417-425, 1987.

(Guy Barles)
Laboratoire de Mathématiques et Physique Théorique
University of Tours
Parc de Grandmont
37200 TOURS, France
E-mail address: barles@univ-tours.fr
URL: http://www.phys.univ-tours.fr/~barles

(Espen R. Jakobsen)
Department of Mathematical Sciences
Norwegian University of Science and Technology
N–7491 Trondheim, Norway
E-mail address: erj@math.ntnu.no
URL: http://www.math.ntnu.no/~erj